

Course: MMDS

Lab 01

Set up Hadoop

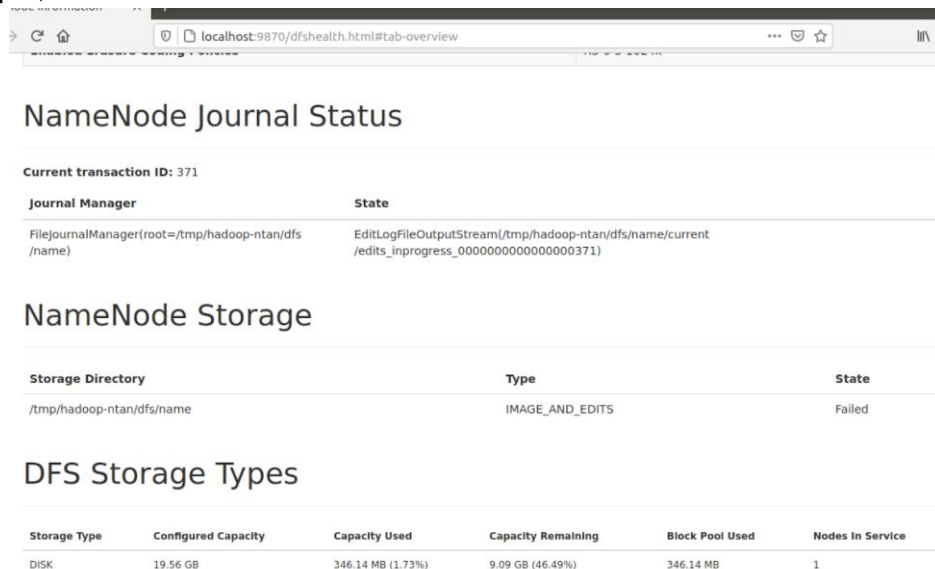
Fill answers of the questions below in the given tables.

Your screenshots must contain commands for required operations.

Question 1:

Set up Hadoop following this [tutorial](#) and then take a screenshot of the hdfs-site ([localhost:9870](#)) which shows information for NameNode Journal Status, NameNode Storage, DFS Storage Types (Overview page).

For example,



The screenshot shows the Hadoop HDFS site at localhost:9870. It displays three sections: NameNode Journal Status, NameNode Storage, and DFS Storage Types. The NameNode Journal Status section shows the current transaction ID as 371 and the journal manager as FileJournalManager. The NameNode Storage section shows the storage directory as /tmp/hadoop-ntan/dfs/name and the type as IMAGE_AND_EDITS. The DFS Storage Types section shows the storage type as DISK, configured capacity as 19.56 GB, capacity used as 346.14 MB (1.73%), capacity remaining as 9.09 GB (46.49%), block pool used as 346.14 MB, and nodes in service as 1.

NameNode Journal Status	
Current transaction ID: 371	
Journal Manager	State
FileJournalManager(root=/tmp/hadoop-ntan/dfs/name)	EditLogFileOutputStream(/tmp/hadoop-ntan/dfs/name/current/edits_inprogress_00000000000000000371)

NameNode Storage		
Storage Directory	Type	State
/tmp/hadoop-ntan/dfs/name	IMAGE_AND_EDITS	Failed

DFS Storage Types					
Storage Type	Configured Capacity	Capacity Used	Capacity Remaining	Block Pool Used	Nodes In Service
DISK	19.56 GB	346.14 MB (1.73%)	9.09 GB (46.49%)	346.14 MB	1

My screenshot:

Activities Firefox Web Browser Mon 10:08

localhost:9870/dfshealth.html#tab-overview

NameNode Journal Status

Current transaction ID: 1

Journal Manager	State
FileJournalManager(root=/tmp/hadoop-pkucpkam/dfs/name)	EditLogFileOutputStream(/tmp/hadoop-pkucpkam/dfs/name/current/edits_inprogress_00000000000000000001)

NameNode Storage

Storage Directory	Type	State
/tmp/hadoop-pkucpkam/dfs/name	IMAGE_AND_EDITS	Active

DFS Storage Types

Storage Type	Configured Capacity	Capacity Used	Capacity Remaining	Block Pool Used	Nodes In Service
DISK	49.45 GB	24 KB (0%)	37.79 GB (76.42%)	24 KB	1

Hadoop, 2019.

Question 2:

Set up Hadoop following this [tutorial](#) and then take a screenshot of the yarn-site ([localhost:8088](#)) which shows information for nodes ([http://localhost:8088/cluster/nodes](#)). For example,

localhost:8088/cluster/nodes

hadoop

Nodes of the cluster

Logged in as: dr:who

Cluster

- About
- Nodes
- Node Labels
- Applications
- NEW
- NEW SAVING
- SUBMITTED
- ACCEPTED
- RUNNING
- FINISHED
- FAILED
- KILLED
- Scheduler

Tools

Cluster Metrics

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Memory Used	Memory Total	Memory Reserved	VCores Used	VCores Total	VCores Reserved
0	0	0	0	0	0 B	8 GB	0 B	0	8	0

Cluster Nodes Metrics

Active Nodes	Decommissioning Nodes	Decommissioned Nodes	Lost Nodes	Unhealthy Nodes	Rebooted Nodes	Shutdown Nodes
1	0	0	0	0	0	0

Scheduler Metrics

Scheduler Type	Scheduling Resource Type	Minimum Allocation	Maximum Allocation	Maximum Cluster Application Priority
Capacity Scheduler	[memory-mb (unit=Mi), vcores]	<memory:1024, vCores:1>	<memory:8192, vCores:4>	0

Show 20 entries
Search:

Node Labels	Rack	Node State	Node Address	Node HTTP Address	Last health-update	Health-report	Containers	Allocation Tags	Mem Used	Mem Avail	VCores Used	VCores Avail	Version
/default-rack	RUNNING	ubuntu:34741	ubuntu:8042	Tue May 11 12:30:58 +0700 2021		0		0 B	8 GB	0	8	3.2.1	

Showing 1 to 1 of 1 entries
First Previous 1 Next Last

My screenshot:

The screenshot shows the Hadoop cluster management interface. The browser window is titled 'Nodes of the cluster' and the address bar shows 'localhost:8088/cluster/nodes'. The page displays the following metrics:

Cluster Metrics

Apps Submitted	Apps Pending	Apps Running	Apps Completed	Containers Running	Memory Used	Memory Total	Memory Reserved	VCoers Used	VCoers Total	VCoers Reserved
0	0	0	0	0	0 B	8 GB	0 B	0	8	0

Nodes of the cluster

Active Nodes	Decommissioning Nodes	Decommissioned Nodes	Lost Nodes	Unhealthy Nodes	Rebooted Nodes	Shutdown Nodes
1	0	0	0	0	0	0

Scheduler Metrics

Scheduler Type	Scheduling Resource Type	Minimum Allocation	Maximum Allocation	Maximum Cluster Application Priority
Capacity Scheduler	[memory.mb (unit=M), vcores]	<memory:1024, vCores:1>	<memory:8192, vCores:4>	0

Nodes of the cluster

Node Labels	Rack	Node State	Node Address	Node HTTP Address	Last health-update	Health-report	Containers	Allocation Tags	Mem Used	Mem Avail	VCoers Used	VCoers Avail	Version
/default-rack		RUNNING	PhucPham.pkucpkam.virtualbox.org:33193	PhucPham.pkucpkam.virtualbox.org:8042	Mon Jan 13 21:20:06 +0700 2025		0		0 B	8 GB	0	8	3.2.1

Showing 1 to 1 of 1 entries